Mair L, Mill AC, Robertson PA, Rushton SP, Shirley MDF, Rodriguez JP, McGowan PJK.

The contribution of scientific research to conservation planning.

Biological Conservation 2018, 223, 82-96.

DOI link to article:

https://doi.org/10.1016/j.biocon.2018.04.037

Date deposited:

10/05/2018

Embargo release date:

09 November 2019

The contribution of scientific research to conservation planning

Abstract

Conservation planning plays an instrumental role in facilitating progress towards biodiversity targets by providing practitioners with the tools required to allocate resources and implement actions. However, the utility of a burgeoning scientific literature to on-the-ground conservation has been questioned. Given such criticisms, and the lack of progress towards the global Aichi Biodiversity Targets, we aim to assess the contribution of scientific research to the field of conservation planning. We applied topic modelling to a body of literature consisting of 4,471 articles pertaining to conservation planning published between 2000-2016. We quantified changes in topic popularity, and assessed the extent to which different topics were addressed within the same articles. We found that research into the status of species and habitats was most prevalent, the process of action planning received considerably less attention, and implementation attracted the least research of all. The scientific literature was thus dominated by biological rather than socio-political research, and furthermore showed a general lack of inter-disciplinary research, which is problematic given that ultimately it is the socio-political context that will determine the success of conservation efforts. The number of publications on implementation and monitoring declined over time, suggesting a waning interest in publishing evidence of plan effectiveness, and that limited efforts have been made to address the 'implementation crisis'. We suggest that filling research gaps, through integration of the social sciences and placing greater value on evidence syntheses, would push scientific research towards greater applicability and help to provide the necessary information to achieve global biodiversity targets.

Keywords

Conservation prioritisation, literature review, research-implementation gap, topic modelling, transdisciplinary research.

1. Introduction

Conservation planning is the process of “deciding where, when and how to allocate limited conservation resources” (Pressey & Bottrill 2009). Planning provides practitioners with the information and direction required to allocate resources and implement actions, ranging from the recovery of endangered species (Clark et al. 2002) to the establishment of large-scale protected area networks (Margules & Pressey 2000). As global conservation targets have evolved since the adoption of the Convention on Biological Diversity (CBD) at the Rio Earth Summit in 1992, the need for conservation planning has become increasingly evident, and planning is now considered essential for achieving the current global Aichi Biodiversity Targets (CBD 2010, 2015). Each of the twenty Aichi targets was designed to contribute towards halting the global loss of biodiversity by 2020 (CBD 2010), and conservation planning should play an instrumental role in facilitating progress towards these targets by providing the strategic framework for the implementation of connected, ecologically representative protected area networks (Aichi Target 11; e.g. Pollock et al. 2017; Venter et al. 2017) and the prevention of species extinctions (Aichi Target 12; e.g. Whitfield et al. 2006; Challender et al. 2014).

Research into conservation planning aims to assist progress towards such ambitious conservation targets, yet there are criticisms about the lack of applicability of much scientific work to practical conservation efforts such as habitat restoration or the designation of protected areas (Knight et al. 2008; Barmuta et al. 2011). Furthermore, current evidence indicates that the majority of the Aichi Biodiversity Targets are unlikely to be met (CBD 2014); species extinctions and declines have not been halted (Tittensor et al. 2014), and while protected area networks are likely to meet the area coverage targets of 17% terrestrial and 10% marine, they do not adequately cover ecoregions or important areas for biodiversity (Butchart et al. 2015). In the context of current conservation shortcomings and deliberations over the utility of research, it is timely to assess the quantity and diversity of scientific research into conservation planning, and hence assess the availability and applicability of information and advice that can build towards achieving global biodiversity targets.

There is a broad range of different conservation planning frameworks outlined in both the scientific and grey literature (Redford et al. 2003; Pressey & Bottrill 2009). These frameworks encompass many steps, with each step falling loosely into three broad categories: (i) assessing the current status of, and threats to, species or areas of conservation interest; (ii) determining what actions should be taken; and (iii) implementation and monitoring (Knight et al. 2006a). The specifics of each step can vary greatly among approaches, and different planning frameworks may diverge on issues such as the process of identifying explicit conservation objectives, and the incorporation of socio-economic considerations (Pressey & Bottrill 2009). Furthermore, planning is a non-linear process, and adaptive management and the revision of plans in response to monitoring outcomes is required for success (Grantham et al. 2010).

The complete conservation planning process is complex, and scientific research projects often focus on in-depth examinations of individual steps or processes within the overall framework. Studies may, for example, assess data requirements (Boitani et al. 2011), incorporate costs estimates (Carwardine et al. 2010), or evaluate the suitability of taxonomic surrogates (Rodrigues & Brooks 2007). This fragmentation of the overall process leads to a large and complex body of literature, and it has been argued that the consideration of individual aspects of the planning process in isolation can result in a disconnect between scientific advance and practical application (Knight et al. 2008). Recent research has suggested that landscape genetics has so far failed to make much impact on conservation planning (Keller et al. 2015) and species distribution models are used less often in planning than might be expected given the proliferation and sophistication of available methods (Tulloch et al. 2016). Furthermore, planning exercises are frequently carried out without the engagement of the end-user or relevant stakeholders, with one review finding that the majority of the publications considered had the aim of improving research techniques rather than achieving implementation (Knight et al. 2008). These issues bring into question the applicability of much of the research pertaining to conservation planning, and emphasise that research direction has different drivers to conservation needs. For example, funding availability has been shown to stimulate research priorities, and this is subject to politics and the changing popularity of research topics (Stroud et al. 2014).

Obtaining an overview of the availability of information in such a vast and complex body of literature is challenging, particularly when the aim is to capture the full extent of the publishing landscape. Topic modelling provides a statistical tool to assess the content of articles in a corpus (a large body of literature; Blei & Lafferty 2009). The approach makes use of the co-occurrence patterns of words in article abstracts to identify a range of topics which represent the main ideas present in a corpus (Griffiths & Steyvers 2004). Topic modelling provides quantitative rigour to summarising themes and allows synthesis across disparate information sources covering different biological, spatial and temporal scales (Westgate et al. 2015). The approach has recently been applied within ecological science to analyse publishing trends in arid ecology research (Greenville et al. 2017), and to compare the topics of conservation-prioritisation articles that did and did not apply species distribution models (Tulloch et al. 2016).

Here, we use topic modelling to assess the contribution of scientific research to the field of conservation planning. We quantify which aspects of the conservation planning process receive the most attention in the published literature, and how topic popularity has changed over time. We also assess the extent to which different aspects of conservation planning are either linked to the broader process or studied in isolation, in order to challenge the implicit assumption that research related to conservation planning is suitable for practical application. We aim to capture the full extent of the publishing landscape; the corpus we analyse consists of 4,471 articles published from 2000-2016 pertaining to conservation planning. Consideration of this large body of literature allows us to determine potential gaps and neglected fields which could be addressed in order to aid progress towards global biodiversity targets.

2. Methods

2.1 Literature search

We searched Web of Science for articles published from 2000-2016 using the terms "conservation plan*" or "recovery plan*" and also "biodiversity", "species", "habitat*" or "ecosystem*". We included only articles published in English and which were categorised as articles or reviews according to document type, giving 4,619 documents.

Citations and abstracts were downloaded and imported in to the program R (R Core Team, 2017) using the package *bibliometrix* (Aria & Cuccurullo 2016). Articles which were categorised as 'in proceedings' and articles that did not have abstracts were removed. This gave 4,471 documents.

2.2 Abstract cleaning

A small number of documents included abstracts written in both English and either Spanish or French; the identification and removal of non-English text is detailed in Appendix A. Abstracts were then transformed into a corpus and processed using the R package *tm* (Feinerer et al. 2008). Search terms were removed as these words were common to all abstracts. Numbers written as words and digits were also removed (Grun & Hornik 2011). The pre-defined list of English stop-words provided in the *tm* package (Feinerer et al. 2008) were removed and we expanded this list by removing the components of abbreviated words on the stop-words list, as well as "also" which was the most common synonym of the stop-word "and" (see Table A1 for list of words removed). Finally, terms added by the publishers for copyright reasons were removed, hyphens and forward slashes were changed to spaces, and all other punctuation was removed (sensu Grun & Hornik 2011).

The suffixes of the abstract words were then removed to reduce words to their common root, and words that appeared in five or fewer articles were removed (following methods in Griffiths & Steyvers 2004; Lu et al. 2017 demonstrated that removal of infrequent words had little impact on model performance). This gave a final corpus with a vocabulary of 4,201 words.

2.3 Topic modelling

Topic modelling defines topics within a corpus based on sets of words that co-occur with unusual frequency (Griffiths & Steyvers 2004; Grun & Hornik 2011). Each topic can be understood as a meaningful combination of ideas within the corpus. Documents belong simultaneously to several topics, making topic modelling an appropriate tool to examine the cross-cutting nature of many research documents.

The inputs to the topic model are a matrix of document-word frequencies and the number of topics to be identified. The most appropriate number of topics for the corpus can determined a priori by carrying out block-cross validation and measuring perplexity (which is a measure of likelihood; Grun & Hornik 2011). The model then provides the weight that each word contributes to a topic, allowing the main ideas of each topic to be inferred. Topic distributions vary over documents and the weight of each topic within a document is provided, which allows the main topic and diversity of topics within a document to be identified.

We identified 40 topics in the corpus by fitting a Latent Dirichlet Allocation (LDA) model with Gibbs sampling using the R package *topicmodels* (Grun & Hornik 2011). Our results from block-cross validation (described in supporting information) indicated that model perplexity decreased as the number of topics modelled increased, suggesting that there were >100 topics in the corpus (see Fig A1). We therefore fitted a model with the number of topics set to 40 in order to balance the need to capture the complexity of the corpus with the need to be able to interpret and communicate results clearly (*sensu* Westgate et al. 2015).

We inspected the 20 highest weighted words in each topic, and considered these alongside a measure of topic similarity, in order to name the topics identified and categorise them into broad themes (*sensu* Westgate et al. 2015; Greenville et al. 2017). Naming and categorisation was done in order to make the presentation of results clearer and more concise. Topic similarity was calculated using the weight that each word contributes to a topic, following methods in Westgate et al. (2015). Each topic was named and then assigned to one of five broad themes, the first two themes were: (i) 'Biome', reflecting that the topic represented a biome or taxonomic group, and (ii) 'Contextual', indicating that the words within the topic provided external context rather than representing a particular aspect of the conservation planning process. The remaining three themes pertained to very broad stages in the conservation planning process: (iii) 'Status Review' includes developing an ecological understanding and related methodologies; (iv) 'Action Planning' includes approaches to determining actions and topics related to the actions themselves; and (v) 'Implementation', which includes monitoring and socio-political considerations.

2.3.1 Topic generality/specificity

Some topics may be general and reflect broad ideas common to many documents within the corpus, while others are more specific. To assess the generality versus specificity of topics, we used the distribution of topic weights within documents. For each document, we selected the topic that received the highest weight. We then calculated the mean weight of a topic when it was selected, and the mean weight of a topic when it was not selected. Plotting these values against each other for all topics gave a comparison of generality versus specificity (Westgate et al. 2015).

2.3.2 Topic popularity

Each document was assigned to the topic that received the highest weight for that document. We then assessed topic popularity based on two metrics. The first was the total number of documents published on each topic during the study period. The second was the change in number of documents published per topic over the course of this period, which indicates whether the prevalence of that topic in the literature is temporally consistent, or shows an increase or decline over time. To do this we fitted a GLMM specified with a Poisson distribution and log-link using the package *lme4* (Bates et al. 2015). The number of documents per topic per year was the response variable, and year and topic were explanatory variables. For each topic, a positive random intercept indicated a higher than average number of documents published on that topic during the study period, and a positive slope indicated an increase in the number of documents published on that topic over time (Westgate et al. 2015). Consideration of these two metrics together allows, for example, identification of 'hot' topics which had a large and increasing number of publications, and 'cold' topics which had a small and decreasing number of publications (Westgate et al. 2015).

2.3.3 Co-occurrence of topics within articles

The distribution of topic weights within documents was used to identify pairs of topics that co-occur within documents. We $\log_{10}$ transformed the matrix of weights with which each document was assigned to each of the 40 topics, and then calculated Euclidean distances. Distances were then scaled from zero to one, where zero indicated that a pair of topics never co-occurred within the same document, and one indicated that a pair of topics always co-occurred in the same document.

2.3.4 The contribution of individual journals

In order to assess the publication contribution of journals to conservation planning, we first quantified the number of documents per journal in the corpus. We then selected the top five journals in terms of their publication volume in this study, and compared the distribution of documents among topics for these journals.

3. Results

We used the 20 highest weighted words per topic and topic similarity (see Table A2 and Fig A3) to name topics and assign them to broad themes (Table 1). The majority of documents in the corpus fell primarily within the theme of Status Review; this included the most frequent topic, which was *Genetics* (Fig 1a). Within the same theme, *Distribution modelling* and *Climate change* were the next most frequent topics. The theme of Action Planning included six topics and overall fewer documents, with *Systematic reserve planning* and *Cost-benefit prioritisation* being the most popular topics. There were only two topics within the theme Implementation; of these *Socio-political considerations* had a higher frequency than *Implementation and monitoring*. Six biomes/taxonomic groups were found to have high enough prevalence in the corpus to be identified as distinct topics and this theme also included a single geographic area, *North America*. Within this theme, the *Marine* topic also included marine protected areas (MPAs), and so this topic had some overlap with the Action Planning theme. The final theme was Contextual and these topics had low frequency. Contextual topics included words that provided the external context of article document, for example whether it described quantitative effects or discussed the current state of research.

3.1 Topic generality/specificity

Topic frequency should be considered alongside topic generality/specificity; the more specific a topic, the more likely it is to be the sole focus (and hence the highest weighted topic) of a document. We found that *Genetics* was a highly specific topic (topic 15; Fig 1b). In contrast, the majority of topics within Action Planning were general (*Decision making, Valuation metrics and indices, Protected areas,* and *Approaches and frameworks*; topics 21, 38, 13 and 33 respectively), which indicated that these topics were broader and therefore often discussed in association with other topics. The Biome topics showed fairly high specificity, with the exception of *North America*, which was intuitive given that this is a geographic area rather than a specific biome or taxon. As would be expected, the Contextual topics tended towards generality.

3.2 Topic popularity

The analysis of topic popularity allowed us to consider the temporal dimension of the publishing landscape. The majority of topics clustered around a slope of zero, indicating relatively small changes in popularity over time (Fig 1c). This included *Genetics*, which had a relatively low slope but large intercept, demonstrating a consistently large number of documents on the topic over time. In contrast, *Climate change* and the *Marine* environment were rapidly expanding 'hot' topics (Fig 1c). Several topics clustered together, with a consistently large number of publications, including *Socio-political considerations* (Fig 1c); the topic which includes 'stakeholder' within the twenty highest weighted words. In contrast, *Implementation and monitoring* and *Systematic reserve selection* had a large number of publications but showed a decline over time, while *North America* could be considered a 'cold' topic as it had both a small and declining number of publications (Fig 1c).

3.3 Co-occurrence of topics

We excluded Contextual topics from the presentation of the analysis of topic co-occurrence because (i) all these topics occur very frequently with topics in other themes and so including them provides little information, and (ii) we were primarily interested in the co-occurrence of topics relating to different parts of the planning process. The correlation matrix showed some expected trends (Fig 2). Within themes, pairs of frequently co-occurring topics included *Community ecology* and *Biological diversity*, *Life history* and *Population ecology*, and *Decision making* and *Approaches and frameworks*. There was also evidence for the co-occurrence of topics from different themes; *Understanding human impacts* (within Status Review) was often considered alongside topics within Action Planning and Implementation. Similarly, *Decision making* (within Action Planning) often co-occurred with topics in Implementation.

The analysis also identified topics that seldom appeared alongside other topics within documents. *Genetics* not only showed relatively low co-occurrence with other Status Review topics, but rarely co-occurred with any topics within Action Planning or Implementation (Fig 2). Similarly, *Life history* was rarely considered alongside Action Planning or Implementation themed topics. This suggests that the research conducted on these two topics is infrequently linked to the later stages in conservation planning.

A perhaps surprising gap was that *Distribution modelling* rarely co-occurred with *Systematic reserve selection* or *Cost-benefit/Prioritisation*, particularly as all three tended to be considered alongside more methodological topics such as *Spatial scale* and *Data collection and use* (Fig 2). *Climate change* was also rarely associated with *Systematic reserve selection* but did show some association with *Decision making* and *Approaches and frameworks*.

3.4 The contribution of individual journals

The corpus consisted of documents from 613 journals, however 50% of documents were contributed by only 25 journals (Fig A4). The distribution of documents among topics in the top five journals (in terms of number of documents) varied among journals and deviated from the overall distribution of topics within the corpus. Biological Conservation had the largest number of documents and, relative to the overall corpus, proportionally more of these were focussed on topics within the Action Planning theme, with a particular emphasis on *Systematic Reserve Selection* (Fig A5a). Conservation Biology was second and similarly showed a high representation of topics within Action Planning, but also a much larger proportion of documents within the Implementation theme than was found overall across the corpus (Fig A5b). More than 10% of documents from Plos One (third in terms of total number of documents) were focussed on *Marine and MPAs* (Fig A5c), while more than 10% of documents from Biodiversity and Conservation (fourth) were on *Biological diversity* (Fig A5d). The topics with the largest number of documents in Diversity and Distributions (fifth) were *Distribution modelling* and *Climate Change*, and this journal had very few documents focussing on Implementation (Fig A5e).

4. Discussion

Although extensive and diverse, we found the scientific literature on conservation planning was dominated by biological rather than socio-political research. Research into the status of species and habitats (primarily biological areas) was the most prevalent, with the process of action planning receiving considerably less attention, while implementation (which requires greater consideration of socio-political considerations) was by far the part of the planning process that attracted the least research. The complete planning process is directional; evidence from Status Review feeds into Action Planning, and both of these stages feed into the outcome of Implementation. Our results show that, overall, publication volume decreases as planning stage progresses, although individual journals (including Biological Conservation and Conservation Biology) evidently do place value on Action Planning and Implementation studies. We suggest that the overall trend could be due to increasing difficulty in achieving publications in later planning stages, which may be in part due to the time lag from plan initiation to implementation. There is also a difference in thematic interest between academics and practitioners due to different drivers and motivations (Habel et al. 2013), and information gained during practice cannot always be translated into scientific publications (Sunderland et al. 2009). We were also able to demonstrate low interconnection among the different parts of the planning process. Many topics within Status Review were rarely considered in the same articles as topics within Action Planning or Implementation, indicating that few articles bridged planning stages and instead tended to take a relatively narrow research focus.

The dominance of biological rather than socio-political research could be seen across the broad planning stages. The Status Review stage consisted of evidence-generating topics associated with developing an ecological understanding (e.g. *Life history* and *Community ecology*) and related methodologies (e.g. *Distribution modelling* and *Spatial analysis*), and was thus concerned primarily with biological analyses. Within the Action Planning theme, *Systematic reserve selection* had the weakest association with *Socio-political considerations* (an Implementation topic) but the largest number of publications, while *Decision-making* had the strongest association but fewest publications. Our results therefore provide quantitative evidence for the long-standing perceptions of both natural and social scientists that biological analyses dominate conservation activities (Fox et al. 2006). This imbalance is problematic as it is the socio-political context that will ultimately determine the success or failure of a conservation project (Balmford & Cowling 2006). Cultural context significantly influences conservation outcomes (Waylen et al. 2010), and a lack of understanding of the social context can result in conservation planning exercises recommending inappropriate or counter-productive actions (Van Vleet et al. 2016). Our results therefore lead us to suggest that conservation planning research would benefit from a greater emphasis on and contribution from social sciences. This conclusion echoes long-standing calls for better integration of the social sciences into conservation planning (e.g. Balmford & Cowling 2006; Pollnac et al. 2010; Christie 2011; Laurance et al. 2012; Ban et al. 2013), which our study demonstrates remain unheeded.

Social sciences play a particularly important role during implementation, which we found to be the planning stage that attracted the fewest publications. Moreover, we found that the number of publications relating primarily to implementation and monitoring has shown a decline over time. This decline is somewhat surprising given that addressing the 'implementation gap', and improving the applicability of research, has long been a major concern in conservation planning (Knight et al. 2008). The need to build an evidence base for the effectiveness of conservation action is well recognised (Sutherland et al. 2004), yet a recent study of a small sample of the conservation planning literature identified a continued lack of reporting of plan implementation and outcomes (Wiersma & Sleep 2016). Our much larger sample of the literature substantiates these conclusions and provides evidence of a declining, rather than growing, interest in this field. While this lack of reporting of implementation may be in part due to limited resources and lack of obligation to report, it may also be due to the absence of a standardised protocol for evaluating the impact of conservation planning processes (McIntosh et al. 2017). Nevertheless, in the absence of such protocols, synthesis techniques such as systematic reviews can be used to bring together evidence from disparate studies (Pullin et al. 2009), and we suggest that placing greater value on such work may help to reverse the apparent decline in publications focussed on implementation, which are much needed to provide the evidence to support conservation policy and management decisions (Sutherland et al. 2004).

Our results also showed weak links between socio-economic and biological topics, indicating a general lack of inter-disciplinary research. For example, a considerable proportion of topics in Status Review (consisting primarily of biological topics) were seldom addressed alongside topics in Action Planning or Implementation (stages which involve greater socio-political considerations). Conservation was defined more than three decades ago as multi-disciplinary and dependant on both biological and social sciences (Soulé 1985) yet our results indicate that, in terms of conducting transdisciplinary research, the gap between biological and social sciences in conservation has rarely been bridged (Fazey et al. 2005a).

A particularly strong example of research which rarely bridges disciplines is the topic of *Genetics*. This is a highly specific evidence-generating topic within Status Review, which has a high publishing volume but the lowest association with outcome topics in later planning stages. It has been suggested that genetic data is gathered because of the relative ease of DNA extraction and analysis, rather than because of a demand from conservation planners (Stinchcombe et al. 2002). Indeed, genetic studies are often considered to have low applicability to conservation, and considerable improvements in design and approach are required to make them more useful (Keller et al. 2015; Shafer et al. 2015). Britt et al. (2018) even go so far as to suggest that in many cases authors of genetic research use conservation to frame their work in order to fit journal specifications, and so rarely offer actionable conservation recommendations. This explanation of ‘framing’ research certainly fits well with our findings, which reflect that in many cases abstracts were highly specific to the topic of genetics, implying that no other topics were relevant enough to merit meaningful inclusion in the abstract. Our results suggest that such framing probably occurs to varying degrees across topics; thus while such research contributes to increased knowledge and understanding, it's unlikely to bridge the 'implementation gap' (Britt et al. 2018). One suggested mechanism to improve the practical applicability of conservation planning research, is to design and execute research in collaboration with both practitioners and relevant stakeholders (Knight et al. 2006b; Ban et al. 2013). This level of engagement is emphasised as essential in the IUCN's conservation planning framework (IUCN/SSC 2008) and greater adoption within the research community would place work naturally within the broader conservation planning framework, resulting in outputs with clear utility.

An important point that emerges from inspection of the topics present in the literature is that the perceived size of the 'implementation gap' may be increased due to the inconsistent use and understanding of the expressions *prioritisation* versus *planning*. Our results show that prioritisation is one step within the broader conservation planning process; we identified *Cost-benefit/prioritisation* as a topic, which we placed within the Action Planning stage. In reviewing the literature, however, we detected some confusion around the utility of prioritisation studies. Spatial prioritisation (e.g. Whitehead et al. 2014), species prioritisation (e.g. Joseph et al. 2009), or prioritisation of actions (e.g. Wilson et al. 2011), may appear to be – or even be specifically presented as – a conservation plan, but in fact each is an exercise within the holistic planning process (Knight et al. 2011). Game et al. (2013) state that "plans are prioritisations"; and while we agree that plans *include* prioritisations, we stress that the two expressions are not synonyms, and prioritisations are *not* plans (McIntosh et al. 2017). The perception that these terms are inter-changeable probably contributes to the perceived implementation gap; since prioritisation exercises alone are not complete plans, they are unlikely to lead directly to conservation action, meaning that much prioritisation work inevitably fails to reach implementation (Kim et al. 2016).

The topic modelling approach we applied here allowed us to quantify and describe the scientific literature on conservation planning. Our results support previous work showing that conservation research in general is taxonomically and geographically biased (Di Marco et al. 2017). We identified a limited number of taxa and biomes, and the presence of only one geographic region, North America, which has previously been shown to make a disproportionately large contribution to the conservation literature (Fazey et al. 2005b). There are, of course, limitations to our approach. For example, the identification of a limited number of biomes and taxa does not mean that other aspects were completely unstudied, rather it reflects the fact that only a few biomes and taxa were studied in sufficient volume to be detected in our analysis. Furthermore, by considering only the article abstracts, the results of our analysis were dependent upon the authors' perceptions and presentation of the content and context of their own work. Nevertheless, the ability to identify topics across several thousand articles makes topic modelling an extremely useful tool for synthesising research (Westgate et al. 2015).

5. Conclusions

The scientific literature on conservation planning is extensive, but the focus of research is not necessarily responding to the needs of those implementing conservation plans. Despite the long-standing acknowledgement that conservation is a multi-disciplinary field (Soulé 1985) and persisting calls for transdisciplinary work (Reyers et al. 2010), we found that research continues to be conducted primarily within unidisciplinary, biological realms. Thus it seems that little has changed in the decade since Knight et al. (2008) argued that fragmentation of the planning process hinders the application of much research to on-the-ground progress. On this basis, we suggest that to increase utility, individual studies should be placed more firmly within the holistic conservation planning framework. As a community, we are not currently learning lessons from the many conservation plans being implemented but not reported in scientific publications. The ultimate causes of species declines and habitat destruction are socio-political (Ceballos et al. 2015), and it is the behaviour and actions of humans that will determine conservation outcomes (Balmford & Cowling 2006). Therefore the biological and socio-political components of conservation problems and solutions need to be considered together to prevent further species extinctions and scale up habitat protection in order to achieve Aichi Biodiversity Targets 11 and 12 (CBD 2010). Consolidation of the different aspects of conservation planning research would result in greater socio-economic integration, push scientific research towards greater applicability, and provide the information and advice required to best inform decisions that will achieve global biodiversity targets.

Literature cited

Aria M, Cuccurullo C. 2016. bibliometrix: a R tool for comprehensive bibliometric analysis of scientific literature. University of Naples Federico II, Naples, Italy.

Balmford A, Cowling RM. 2006. Fusion or Failure? The Future of Conservation Biology. Conservation Biology **20**:692-695.

Ban NC, et al. 2013. A social-ecological approach to conservation planning: embedding social considerations. Frontiers in Ecology and the Environment **11**:194-202.

Barmuta LA, Linke S, Turak E. 2011. Bridging the gap between ‘planning’ and ‘doing’ for biodiversity conservation in freshwaters. Freshwater Biology **56**:180-195.

Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software **67**:1-48.

Blei DM, Lafferty JD. 2009. Topic Models in Srivastava A, and Sahami M, editors. Text Mining: Classification, Clustering and Applications. Chapman & Hall/CRC Press.

Boitani L, Maiorano L, Baisero D, Falcucci A, Visconti P, Rondinini C. 2011. What spatial data do we need to develop global mammal conservation strategies? Philosophical Transactions of the Royal Society B: Biological Sciences **366**:2623.

Britt M, Haworth SE, Johnson JB, Martchenko D, Shafer ABA. 2018. The importance of non-academic coauthors in bridging the conservation genetics gap. Biological Conservation **218**:118-123.

Butchart SHM, et al. 2015. Shortfalls and Solutions for Meeting National and Global Conservation Area Targets. Conservation Letters **8**:329-337.

Carwardine J, Wilson KA, Hajkowicz SA, Smith RJ, Klein CJ, Watts M, Possingham HP. 2010. Conservation Planning when Costs Are Uncertain. Conservation Biology **24**:1529-1537.

CBD. 2010. Convention on Biological Diversity: Decision X/2: the strategic plan for biodiversity 2011-2020 and the Aichi biodiversity targets. Nagoya, Japan.

CBD. 2014. Global Biodiversity Outlook 4. CBD, Montréal.

CBD. 2015. Species conservation assessment as an essential element of achieving Aichi Biodiversity Target 12. Montreal, Canada.

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. 2015. Accelerated modern human–induced species losses: Entering the sixth mass extinction. Science Advances **1**.

Challender DWS, Waterman C, Baillie JEM. 2014. Scaling up pangolin conservation. IUCN SSC Pangolin Specialist Group Conservation Action Plan. . Zoological Society of London. London, UK.

Christie P. 2011. Creating space for interdisciplinary marine and coastal research: five dilemmas and suggested resolutions. Environmental Conservation **38**:172-186.

Clark JA, Hoekstra JM, Boersma PD, Kareiva P. 2002. Improving U.S. Endangered Species Act Recovery Plans: Key Findings and Recommendations of the SCB Recovery Plan Project. Conservation Biology **16**:1510-1519.

Di Marco M, et al. 2017. Changing trends and persisting biases in three decades of conservation science. Global Ecology and Conservation **10**:32-42.

Fazey I, Fischer J, Lindenmayer DB. 2005a. What do conservation biologists publish? Biological Conservation **124**:63-73.

Fazey I, Fischer J, Lindenmayer DB. 2005b. Who does all the research in conservation biology? Biodiversity & Conservation **14**:917-934.

Feinerer I, Hornik K, Meyer D. 2008. Text Mining Infrastructure in R. Journal of Statistical Software; Vol 1, Issue 5 (2008).

Fox HE, Christian C, Nordby JC, Pergams ORW, Peterson GD, Pyke CR. 2006. Perceived Barriers to Integrating Social Science and Conservation. Conservation Biology **20**:1817-1820.

Game ET, Kareiva P, Possingham HP. 2013. Six Common Mistakes in Conservation Priority Setting. Conservation Biology **27**:480-485.

Grantham HS, Bode M, McDonald-Madden E, Game ET, Knight AT, Possingham HP. 2010. Effective conservation planning requires learning and adaptation. Frontiers in Ecology and the Environment **8**:431-437.

Greenville AC, Dickman CR, Wardle GM. 2017. 75 years of dryland science: Trends and gaps in arid ecology literature. PLOS ONE **12**:e0175014.

Griffiths TL, Steyvers M. 2004. Finding scientific topics. Proceedings of the National Academy of Sciences **101**:5228-5235.

Grun B, Hornik K. 2011. Topicmodels: An R Package for Fitting Topic Models. Journal of Statistical Software **40**:1-30.

Habel JC, Gossner MM, Meyer ST, Eggermont H, Lens L, Dengler J, Weisser WW. 2013. Mind the gaps when using science to address conservation concerns. Biodiversity and Conservation **22**:2413-2427.

IUCN/SSC. 2008. Strategic Planning for Species Conservation: A Handbook. Version 1.0. Gland, Switzerland: IUCN Species Survival Commission.

Joseph LN, Maloney RF, Possingham HP. 2009. Optimal Allocation of Resources among Threatened Species: a Project Prioritization Protocol. Conservation Biology **23**:328-338.

Keller D, Holderegger R, van Strien MJ, Bolliger J. 2015. How to make landscape genetics beneficial for conservation management? Conservation Genetics **16**:503-512.

Kim MK, Evans L, Scherl LM, Marsh H. 2016. The User, not the Tool: Perceptions of Credibility and Relevance Affect the Uptake of Prioritisation. Environmental Management **57**:836-846.

Knight AT, Cowling RM, Boshoff AF, Wilson SL, Pierce SM. 2011. Walking in STEP: Lessons for linking spatial prioritisations to implementation strategies. Biological Conservation **144**:202-211.

Knight AT, Cowling RM, Campbell BM. 2006a. An Operational Model for Implementing Conservation Action. Conservation Biology **20**:408-419.

Knight AT, Cowling RM, Rouget M, Balmford A, Lombard AT, Campbell BM. 2008. Knowing But Not Doing: Selecting Priority Conservation Areas and the Research–Implementation Gap. Conservation Biology **22**:610-617.

Knight AT, et al. 2006b. Designing systematic conservation assessments that promote effective implementation: Best practice from South Africa. Conservation Biology **20**:739-750.

Laurance WF, Koster H, Grooten M, Anderson AB, Zuidema PA, Zwick S, Zagt RJ, Lynam AJ, Linkie M, Anten NPR. 2012. Making conservation research more relevant for conservation practitioners. Biological Conservation **153**:164-168.

Lu K, Cai X, Ajiferuke I, Wolfram D. 2017. Vocabulary size and its effect on topic representation. Information Processing & Management **53**:653-665.

Margules CR, Pressey RL. 2000. Systematic conservation planning. Nature **405**:243-253.

McIntosh EJ, Pressey RL, Lloyd S, Smith RJ, Grenyer R. 2017. The Impact of Systematic Conservation Planning. Annual Review of Environment and Resources **42**:677-697.Pollnac R, Christie P, Cinner JE, Dalton T, Daw TM, Forrester GE, Graham NAJ, McClanahan TR. 2010. Marine reserves as linked social-ecological systems. Proceedings of the National Academy of Sciences of the United States of America **107**:18262-18265.

Pollock LJ, Thuiller W, Jetz W. 2017. Large conservation gains possible for global biodiversity facets. Nature **546**:141-144.

Pressey RL, Bottrill MC. 2009. Approaches to landscape- and seascape-scale conservation planning: convergence, contrasts and challenges. Oryx **43**:464-475.

Pullin AS, Knight TM, Watkinson AR. 2009. Linking reductionist science and holistic policy using systematic reviews: unpacking environmental policy questions to construct an evidence-based framework. Journal of Applied Ecology **46**:970-975.

Redford KH, et al. 2003. Mapping the conservation landscape. Conservation Biology **17**:116-131.

Reyers B, Roux DJ, Cowling RM, Ginsburg AE, Nel JL, Farrell PO. 2010. Conservation Planning as a Transdisciplinary Process. Conservation Biology **24**:957-965.

Rodrigues ASL, Brooks TM. 2007. Shortcuts for biodiversity conservation planning: The effectiveness of surrogates. Annual Review of Ecology Evolution and Systematics **38**:713-737.

Shafer ABA, et al. 2015. Genomics and the challenging translation into conservation practice. Trends in Ecology & Evolution **30**:78-87.

Soulé ME. 1985. What is conservation biology? Bioscience **35**:727-734.

Stinchcombe J, Moyle LC, Hudgens BR, Bloch PL, Chinnadurai S, Morris WF. 2002. The influence of the academic conservation biology literature on endangered species recovery planning. Conservation Ecology **6**.

Stroud JT, Rehm E, Ladd M, Olivas P, Feeley KJ. 2014. Is conservation research money being spent wisely? Changing trends in conservation research priorities. Journal for Nature Conservation **22**:471-473.

Sunderland T, Sunderland-Groves J, Shanley P, Campbell B. 2009. Bridging the Gap: How Can Information Access and Exchange Between Conservation Biologists and Field Practitioners be Improved for Better Conservation Outcomes? Biotropica **41**:549-554.

Sutherland WJ, Pullin AS, Dolman PM, Knight TM. 2004. The need for evidence-based conservation. Trends in Ecology & Evolution **19**:305-308.

Tittensor DP, et al. 2014. A mid-term analysis of progress toward international biodiversity targets. Science **346**:241-244.

Tulloch AIT, Sutcliffe P, Naujokaitis-Lewis I, Tingley R, Brotons L, Ferraz K, Possingham H, Guisan A, Rhodes JR. 2016. Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. Biological Conservation **199**:157-171.

Van Vleet E, Bray DB, Durán E. 2016. Knowing but not knowing: Systematic conservation planning and community conservation in the Sierra Norte of Oaxaca, Mexico. Land Use Policy **59**:504-515.

Venter O, Magrach A, Outram N, Klein CJ, Marco MD, Watson JEM. 2017. Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. Conservation Biology.

Waylen KA, Fischer A, McGowan PJK, Thirgood SJ, Milner-Gulland EJ. 2010. Effect of Local Cultural Context on the Success of Community-Based Conservation Interventions. Conservation Biology **24**:1119-1129.

Westgate MJ, Barton PS, Pierson JC, Lindenmayer DB. 2015. Text analysis tools for identification of emerging topics and research gaps in conservation science. Conservation Biology **29**:1606-1614.

Whitehead AL, Kujala H, Ives CD, Gordon A, Lentini PE, Wintle BA, Nicholson E, Raymond CM. 2014. Integrating Biological and Social Values When Prioritizing Places for Biodiversity Conservation. Conservation Biology **28**:992-1003.

Whitfield DP, Fielding AH, McLeod DRA, Haworth PF, Watson J. 2006. A conservation framework for the golden eagle in Scotland: Refining condition targets and assessment of constraint influences. Biological Conservation **130**:465-480.

Wiersma YF, Sleep DJH. 2016. A review of applications of the six-step method of systematic conservation planning. Forestry Chronicle **92**:322-335.

Wilson KA, Evans MC, Di Marco M, Green DC, Boitani L, Possingham HP, Chiozza F, Rondinini C. 2011. Prioritizing conservation investments for mammal species globally. Philosophical Transactions of the Royal Society B: Biological Sciences **366**:2670.

Tables

Table 1. Topic number, the five highest weighted words, topic name (which was based on the twenty highest weighted words, see Table A2) and theme.

No.	Top 5 words (stemmed)	Topic name	Theme
1	New, research, need, mani, global	State of research	Contextual
2	Threaten, threat, risk, assess, extinct	Threatened status	Status review
3	Bird, wetland, breed, abund, winter	Birds and wetlands	Biome
4	Endem, distribut, south, mammal, hotspot	Endemism and hotspots	Status review
5	Program, monitor, endang, implement, wildlif	Implementation and monitoring	Implementation
6	Community, environment, composit, variabl, assemblag	Community ecology	Status review
7	Use, rang, select, behaviour, anim	Habitat use	Status review
8	River, fish, water, freshwat, stream	Freshwater	Biome
9	Marin, fish, sea, coastal, fisheri	Marine and MPAs	Biome
10	Develop, impact, restor, service, product	Understanding human impacts	Status review
11	Rang, state, north, unit, associ	North America	Biome
12	Chang, climat, future, project, current	Climate change	Status review
13	Area, protect, priority, identify, exist	Protected areas	Action planning
14	Spatial, scale, local, larg, across	Spatial scale	Status review
15	Genet, popul, among, divers, structur	Genetics	Status review
16	Plant, soil, seed, island, nativ	Plants	Biome
17	Effect, factor, influenc, import, interact	Describing effects	Contextual
18	Site, differ, size, compar, signific	Comparative	Contextual
19	Nest, success, year, femal, reproduct	Life history	Status review
20	Use, map, data, method, base	Spatial analysis	Status review

21	Manag, decis, inform, strategi, uncertainti	Decision making	Action planning
22	System, ecology, type, level, biolog	Categorisation	Contextual
23	Studi, result, show, differ, import	Interpretation of results	Contextual
24	Popul, rate, size, growth, individu	Population ecology	Status review
25	Landscap, connect, fragment, patch, dispers	Landscape connectivity	Status review
26	Cost, benefit, object, priority, action	Cost-benefit / prioritisation	Action planning
27	Land, use, cover, landscap, agricultur	Land use	Status review
28	Region, high, area, pattern, elev	Biogeography	Status review
29	Divers, rich, group, indic, taxa	Biological diversity	Status review
30	Human, density, disturb, active, predat	Human-wildlife interactions	Status review
31	Model, predict, distribut, suitabl, variabl	Distribution modelling	Status review
32	Local, social, polici, process, natur	Socio-political considerations	Implementation
33	Approach, process, identify, integr, framework	Approaches and frameworks	Action planning
34	Forest, tree, fire, cover, stand	Forest	Biome
35	Time, term, year, long, increas	Temporal trends	Status review
36	Reserve, select, target, network, design	Systematic reserve selection	Action planning
37	Estim, survey, sampl, abund, detect	Surveys	Status review
38	Use, valu, measure, base, indic	Valuation indices and metrics	Action planning
39	Increase, level, effect, function, high	Quantitative effects	Contextual
40	Data, use, collect, inform, avail	Data collection and use	Status review

Figures

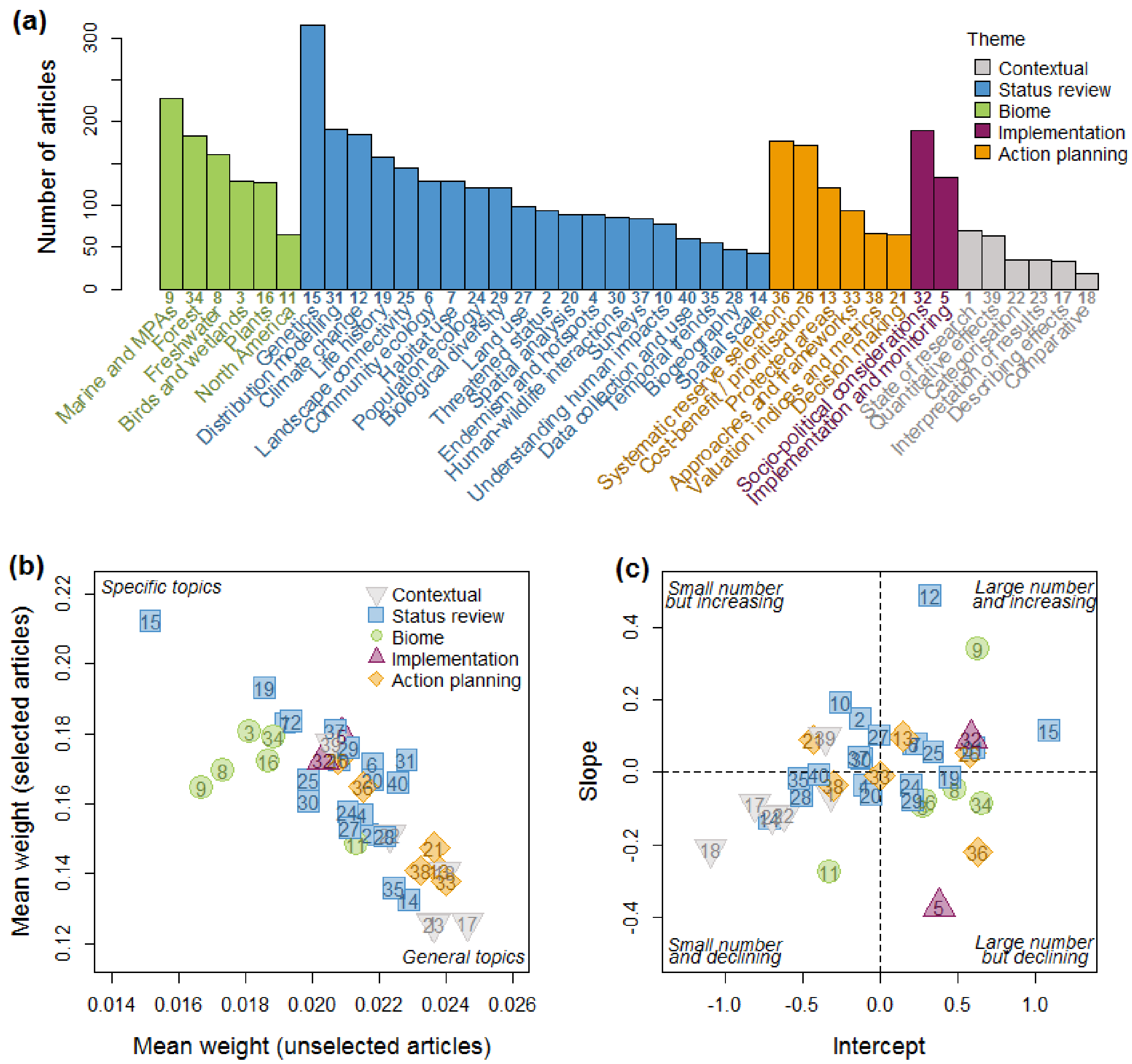


Figure 1. **(a)** Topic frequency in the corpus. Each article was assigned to the topic with the highest weight. The x-axis gives both the topic name and topic number. **(b)** Topic generality/specificity. Topics in the top left hand corner are specific (more likely to be the sole topic present within an article), while topics in the bottom right are general (broad topics common to many articles within the corpus). **(c)** Topic popularity. Topics that have increased in popularity over time have a positive slope, while topics that have declined in popularity have a negative slope (y-axis). Topics with a higher than average number of publications have a positive intercept, and those with a lower than

average number of publications have a negative intercept (x-axis). The top five words associated with each topic can be found in Table 1, alongside the topic themes, which are indicated by the colours and shapes.

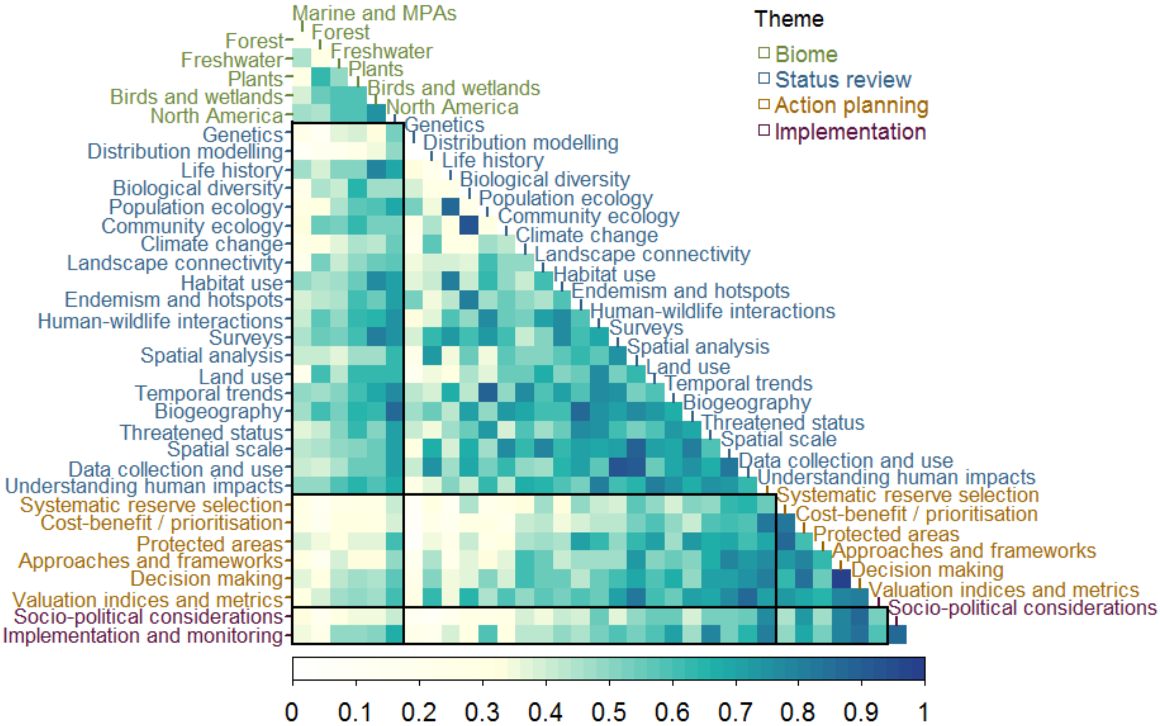


Figure 2. Correlation matrix of topic co-occurrence within articles. Zero indicates that a pair of topics never co-occur within the same article, and one indicates that a pair of topics always co-occur in the same article. Black outlined boxes indicate between-theme comparisons.

Appendix A

Removal of Spanish and French text

Some abstracts were written in both English and either Spanish (n = 39) or French (n = 3). To identify the Spanish abstracts, we created a list of 8 very common Spanish words ("que", "y", "los", "las", "con", "una", "del", "de", "la") and searched the abstracts for them. For abstracts containing at least seven of these words, we split the abstracts based on the first occurrence of one of these words, and maintained only the text occurring before the Spanish word (English always came first). For abstracts also written in French, we created a list of 8 very common French words ("des", "est", "les", "sont", "que", "sur", "avec", "pas", "mais"), and identified French abstracts based on the presence of at least six of these. French abstracts were split using the word “resume”, and only text before this word was maintained.

Removal of stop-words

The words removed from abstracts are given in Table A.1. These include the pre-defined list of English stop-words provided in the *tm* package (Feinerer et al. 2008), the components of abbreviated words on the stop-words list, “also” which was the most common synonym of the stop-word “and”, the search terms, numbers written as words and terms added by the publishers for copyright reasons were removed.

Table A1. Words removed from abstracts during abstract cleaning.

Words removed from abstracts						
Stopwords						
a	about	above	after	again	against	all
am	an	and	any	are	aren't	as
at	be	because	been	before	being	below
between	both	but	by	can't	cannot	could
couldn't	did	didn't	do	does	doesn't	doing
don't	down	during	each	few	for	from
further	had	hadn't	has	hasn't	have	haven't
having	he	he'd	he'll	he's	her	here
here's	hers	herself	him	himself	his	how
how's	i	i'd	i'll	i'm	i've	if
in	into	is	isn't	it	it's	its
itself	let's	me	more	most	mustn't	my
myself	no	nor	not	of	off	on
once	only	or	other	ought	our	ours
ourselves	out	over	own	same	shan't	she
she'd	she'll	she's	should	shouldn't	so	some
such	than	that	that's	the	their	theirs
them	themselves	then	there	there's	these	they
they'd	they'll	they're	they've	this	those	through
to	too	under	until	up	very	was
wasn't	we	we'd	we'll	we're	we've	were
weren't	what	what's	when	when's	where	where's
which	while	who	who's	whom	why	why's
with	won't	would	wouldn't	you	you'd	you'll
you're	you've	your	yours	yourself	yourselves	
Stems of abbreviated words in stopword list						
can	will					
Most common synonym of stopword list						
also						
Search terms						
biodiversity	conservation	ecosystem	habitat	plan	recovery	species
Numbers written as words						
one	two	three	four	five	six	seven
eight	nine	ten				
Terms added to abstracts by publishers						
“all rights reserved”	(c)	copyright	elsevier	john	ltd	wiley

Block cross validation to determine number of topics

We investigated the performance of models with varying numbers of topics using 10-fold block cross-validation. The corpus was randomly divided into ten equal parts, and each part in turn was withheld from the model fitting process. Model performance was then tested by calculating perplexity on the withheld data. Perplexity indicates the uncertainty in predicting a single word; the lower the perplexity value, the better the model performance, and a perplexity equal to the size of the vocabulary indicates a performance no better than chance (Griffiths & Steyvers 2004).

We found that perplexity decreased as the number of topics increased (Fig A1), indicating that the corpus consists of a large number of topics.

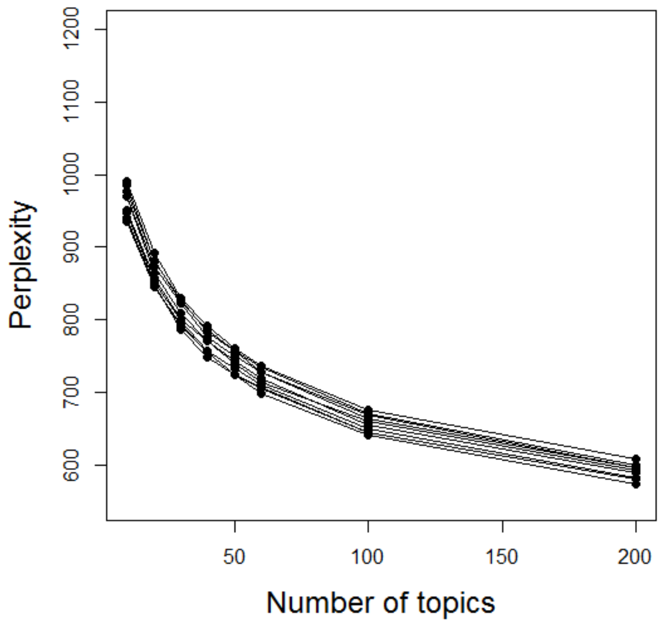


Figure A1. Perplexity against number of topics for 10-fold block cross-validation.

Topic names and highest weighted words

Word weights declined steeply over the first few words (Fig. A2), therefore we used the 20 highest weighted words per topic (Table A1) to determine a name for each topic which succinctly represented the main ideas present in the topic (Table 1 main text). We also used topic similarity (calculated based on word weights within each topic; see main text), to inform topic naming (Fig. A3).

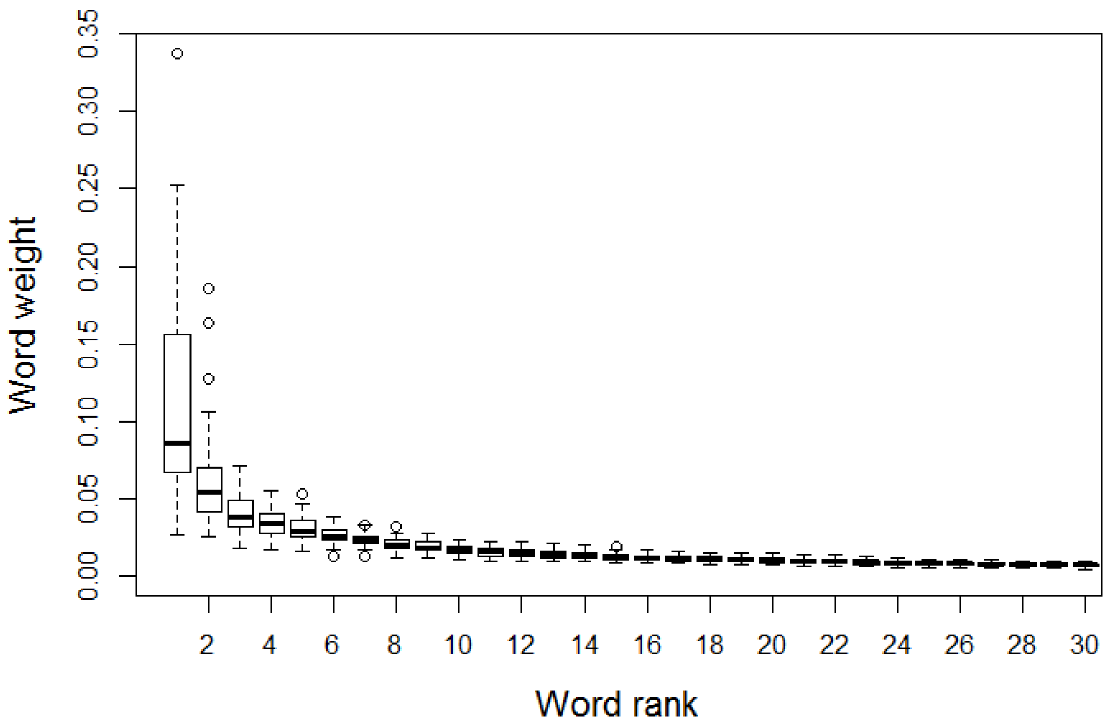


Figure A2. Word weights against word rank for the 30 highest weighted words within each of the 40 topics modelled.

Table A2. The 20 highest weighted words for each topic.

No.	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
1	new	threaten	bird	endem	program	communiti	use	river	marin	develop
2	research	threat	wetland	distribut	monitor	environment	rang	fish	fish	impact
3	need	risk	breed	south	endang	composit	select	water	sea	restor
4	mani	assess	abund	mammal	implement	variabl	behavior	freshwat	coastal	servic
5	global	extinct	winter	hotspot	wildlif	assemblag	anim	stream	fisheri	product
6	world	list	migrat	amphibian	effort	variat	movement	basin	reef	potenti
7	current	loss	use	africa	state	structur	individu	aquat	turtl	provid
8	knowledg	vulner	field	rang	manag	differ	locat	watersh	coral	mitig
9	present	status	season	studi	act	pattern	forag	riparian	ocean	reduc
10	countri	red	import	vertebr	critic	type	avail	lake	stock	effect
11	although	nation	pool	high	design	gradient	studi	reach	mpa	sustain
12	major	critic	avian	region	improv	similar	prefer	catchment	mpas	improv
13	recent	endang	spring	import	action	explain	resourc	salmon	island	need
14	inform	iucn	migratori	biogeograph	goal	relat	home	flow	depth	includ
15	gap	level	increas	geograph	agenc	relationship	track	qualiti	manag	increas
16	part	face	food	identifi	requir	analyzi	space	trout	water	carbon
17	provid	global	ground	signific	recommend	among	bat	restor	estuari	offset
18	particular	addit	warbler	base	develop	veget	activ	within	atlant	environment
19	focus	criteria	studi	fauna	includ	abund	food	condit	coast	natur
20	larg	action	may	biom	review	use	roost	hydrolog	beach	strategi

Table A2 ctd.

No.	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
1	rang	chang	area	spatial	genet	plant	effect	site	nest	use
2	state	climat	protect	scale	popul	soil	factor	differ	success	map
3	north	futur	prioriti	local	among	seed	influenc	size	year	data
4	unit	project	identifi	larg	divers	island	import	compar	femal	method
5	associ	current	exist	across	structur	nativ	interact	signific	reproduct	base
6	southern	rang	park	pattern	evolutionari	rare	affect	higher	adult	inform
7	northern	scenario	nation	region	gene	invas	posit	within	surviv	resolut
8	california	distribut	current	multipl	differenti	veget	relat	among	breed	assess
9	western	impact	network	broad	variat	natur	relationship	suggest	life	provid
10	includ	shift	within	extent	within	insect	associ	number	male	remot
11	within	condit	zone	fine	phylogenet	studi	studi	small	juvenil	techniqu
12	distribut	respons	gap	explicit	distinct	graze	negat	may	age	analyzi
13	current	potenti	outsid	within	isol	establish	suggest	larger	stage	identifi
14	canada	adapt	establish	process	sequenc	butterfli	role	result	studi	sens
15	central	like	system	assess	morpholog	situ	understand	lower	individu	deriv
16	occur	global	addit	import	marker	host	may	similar	mean	tool
17	america	may	pas	complex	diverg	non	characterist	greater	season	combin
18	greater	suitabl	expand	vari	geograph	field	like	found	owl	grid
19	across	result	core	heterogen	flow	collect	result	averag	histori	cell
20	mexico	increas	locat	locat	dna	pollin	strong	whether	spot	area

Table A2 ctd.

No.	Topic 21	Topic 22	Topic 23	Topic 24	Topic 25	Topic 26	Topic 27	Topic 28	Topic 29	Topic 30
1	manag	system	studi	popul	landscap	cost	land	region	divers	human
2	decis	ecolog	result	rate	connect	benefit	use	high	rich	densiti
3	inform	type	show	size	fragment	object	cover	area	group	disturb
4	strategi	level	differ	growth	patch	prioriti	landscap	pattern	indic	activ
5	uncertainti	biolog	import	individu	dispers	action	agricultur	elev	taxa	predat
6	make	use	main	declin	distanc	scenario	urban	mountain	surrog	anthropogen
7	resourc	defin	analyz	endang	corridor	effect	develop	studi	taxonom	high
8	provid	unit	aim	mortal	function	achiev	natur	distribut	pattern	road
9	support	classif	analyzi	viabil	matrix	priorit	veget	zone	correl	increas
10	applic	base	first	demograph	structur	econom	area	ecoregion	taxon	pressur
11	tool	natur	european	surviv	maintain	approach	grassland	slope	beetl	prey
12	limit	repres	sever	wild	isol	trade	increas	geograph	invertebr	impact
13	appli	compon	found	small	import	effici	type	cluster	beta	tiger
14	adapt	determin	mediterranean	persist	loss	outcom	convers	china	total	larg
15	approach	class	investig	paramet	movement	opportun	intens	central	bird	potenti
16	action	attribut	order	simul	dynam	invest	buffer	along	high	carnivor
17	effect	concept	consid	transloc	metapopul	alloc	privat	identifi	test	natur
18	altern	geograph	thus	reduc	network	target	surround	low	congruenc	wildlif
19	evalu	basi	suggest	vital	persist	resourc	rural	bear	repres	low
20	improv	contribut	account	estim	qualiti	optim	adjac	import	studi	hunt

Table A2 ctd.

No.	Topic 31	Topic 32	Topic 33	Topic 34	Topic 35	Topic 36	Topic 37	Topic 38	Topic 39	Topic 40
1	model	local	approach	forest	time	reserv	estim	use	increas	data
2	predict	social	process	tree	term	select	survey	valu	level	use
3	distribut	polici	identifi	fire	year	target	sampl	measur	effect	collect
4	suitabl	process	integr	cover	long	network	abund	base	function	inform
5	variabl	natur	framework	stand	increas	design	detect	indic	high	avail
6	use	govern	assess	oak	declin	use	occup	high	low	distribut
7	occurr	resourc	base	tropic	period	represent	use	index	may	set
8	perform	stakehold	ecolog	veget	trend	repres	probabl	metric	condit	bias
9	presenc	knowledg	develop	canopi	dynam	systemat	method	evalu	degre	record
10	potenti	practic	requir	domin	tempor	set	provid	identifi	respons	limit
11	environment	scienc	appli	log	histor	algorithm	conduct	differ	temperatur	dataset
12	predictor	econom	need	plantat	recent	featur	monitor	rank	surfac	common
13	nich	communiti	specif	old	past	effici	densiti	method	trait	howev
14	evalu	scientif	applic	structur	decad	optim	number	prioriti	relat	well
15	base	environment	address	deforest	last	consid	effort	assess	measur	occurr
16	regress	public	propos	larg	sinc	exist	mean	threshold	concentr	often
17	valid	implement	systemat	harvest	short	priorit	result	relat	caus	inventori
18	best	ecolog	incorpor	pine	observ	approach	observ	expert	affect	sourc
19	probabl	valu	multipl	secondari	understand	solut	requir	qualiti	indic	lack
20	sdms	sustain	key	ant	histori	base	total	criteria	sensit	repres

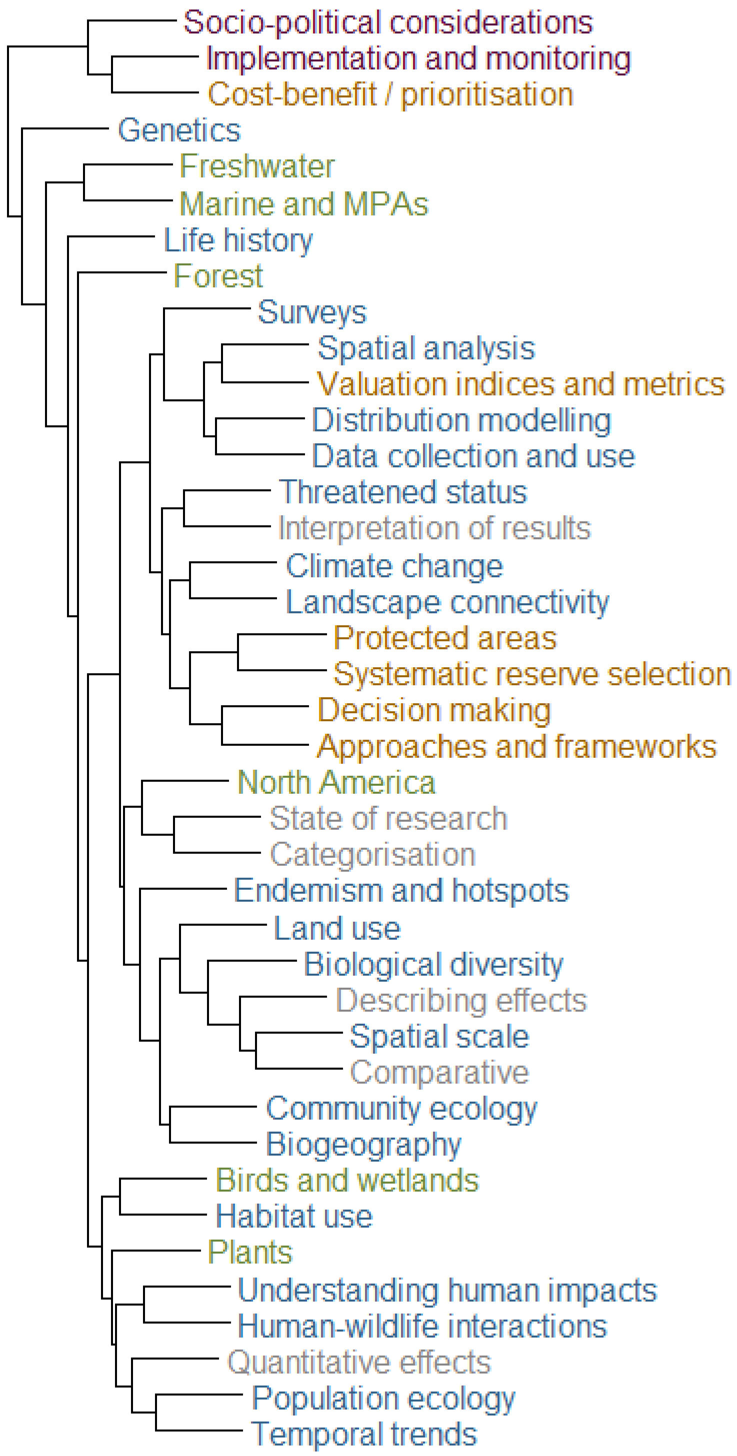


Figure A3. Topic similarity, calculated using the weight that each word contributes to a topic, following methods in Westgate et al. (2015). Topics which are found to be similar based on word weights are grouped more closely in the dendrogram. Colours indicate the topic themes given in Table 1 in the main text.

The contribution of individual journals

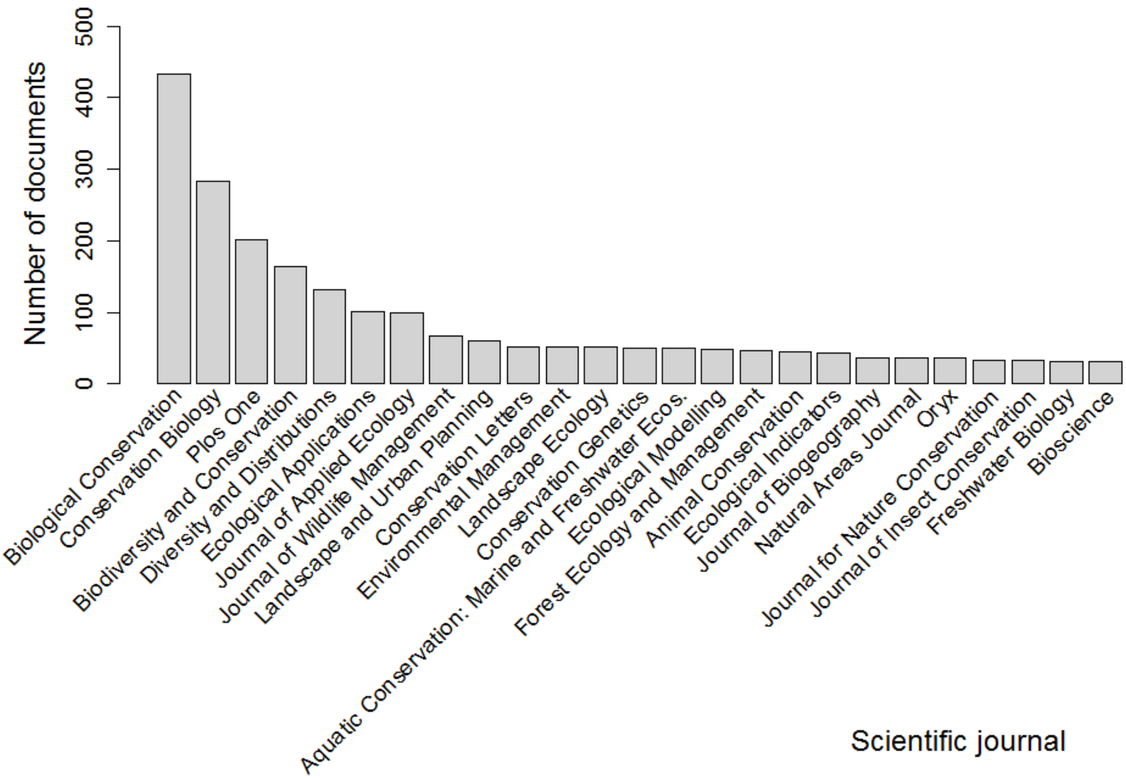
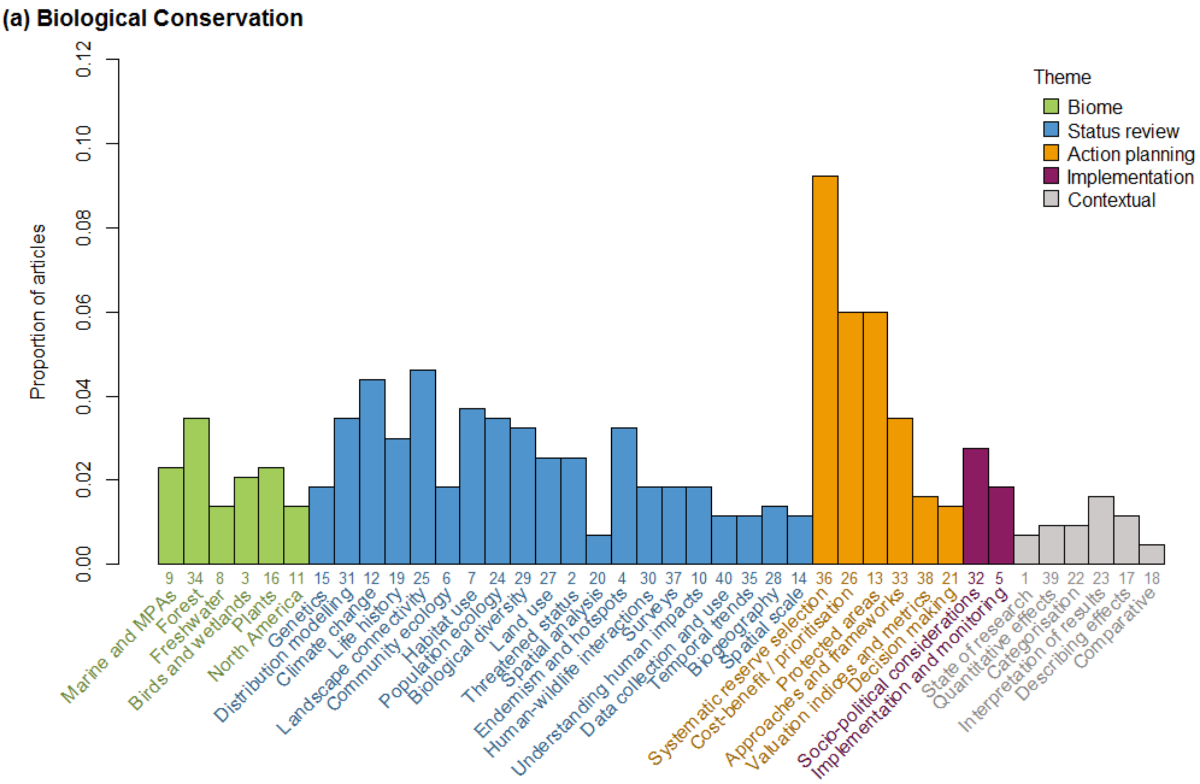
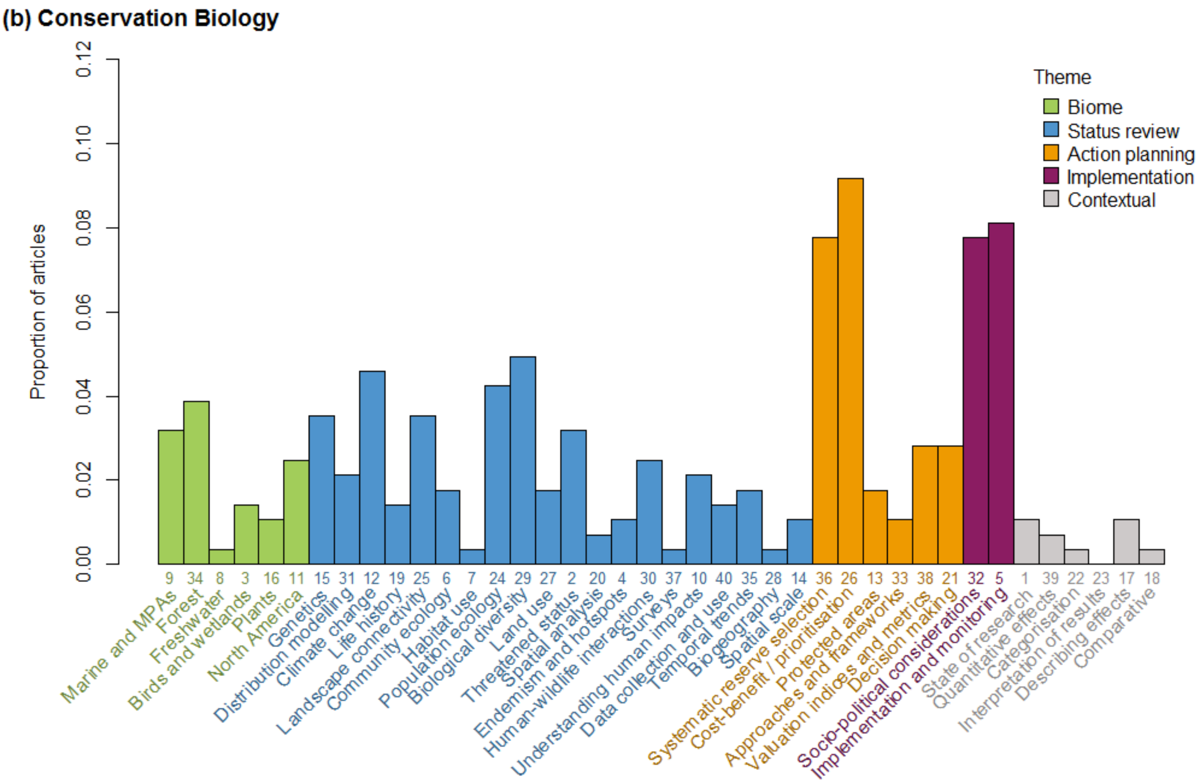


Figure A4. The number of documents contributed to the corpus by each scientific journal, for the 25 journals with largest number of documents.

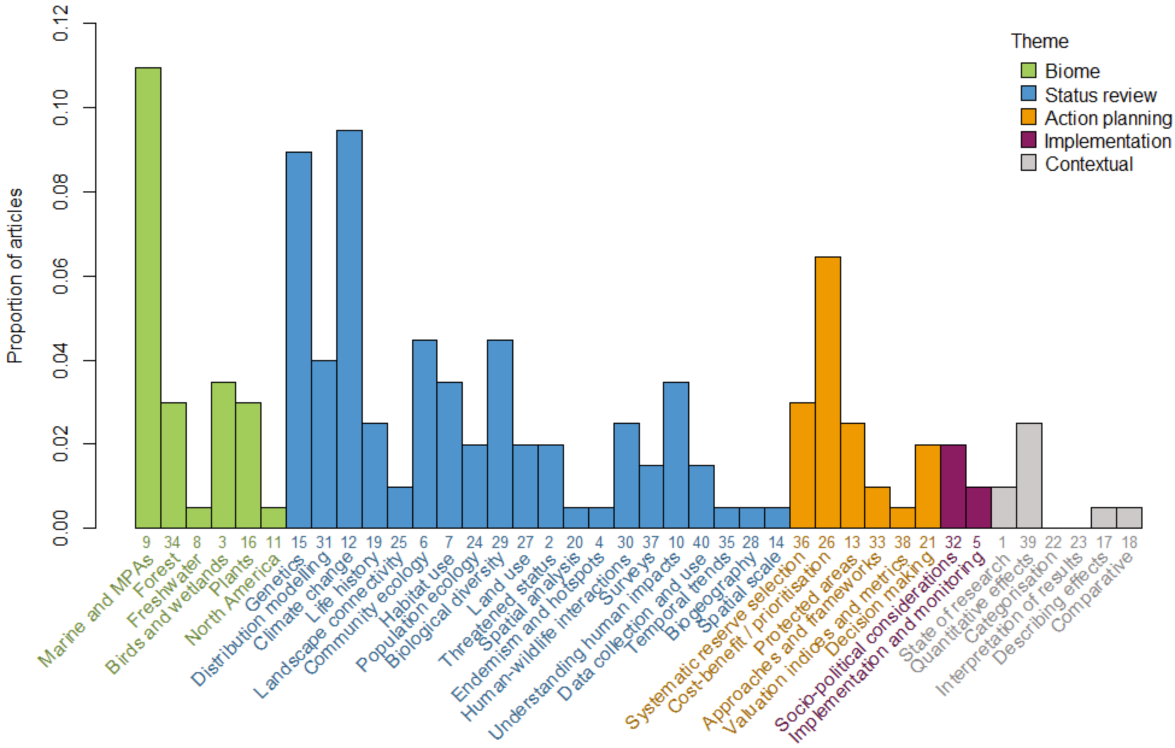


(a) Biological Conservation
Proportion of articles
Theme
Biome
Status review
Action planning
Implementation
Contextual
Marine and MPAs
Forest
Freshwater
Birds and wetlands
Plants
North America
Genetics
Distribution modelling
Climate change
Life history
Landscape connectivity
Community ecology
Habitat use
Population ecology
Biological diversity
Land use
Threatened status
Spatial analysis
Endemism and hotspots
Human-wildlife interactions
Surveys
Understanding human impacts
Data collection and use
Temporal trends
Biogeography
Spatial scale
Systematic reserve selection
Cost-benefit / prioritisation
Protected areas
Approaches and frameworks
Valuation indices and metrics
Decision making
Socio-political considerations
Implementation and monitoring
State of research
Quantitative effects
Categorisation
Interpretation of results
Describing effects
Comparative

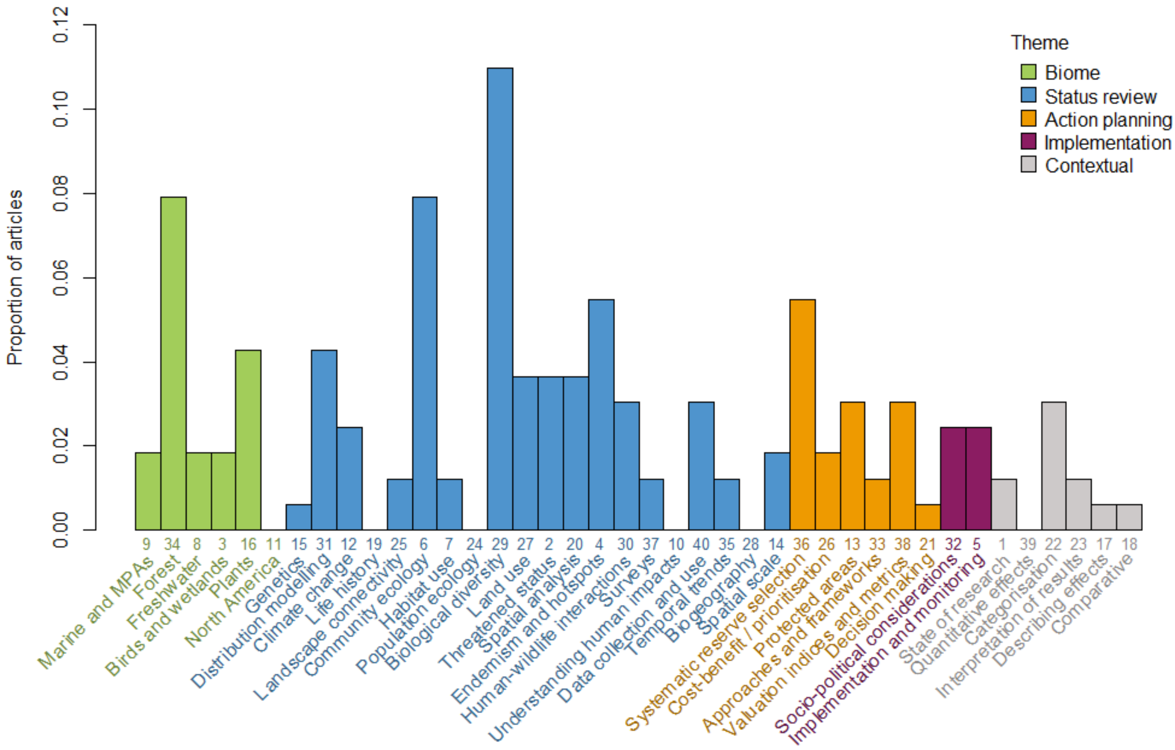


(b) Conservation Biology
Proportion of articles
Theme
Biome
Status review
Action planning
Implementation
Contextual

(c) Plos One



(d) Biodiversity and Conservation



(e) Diversity and Distributions

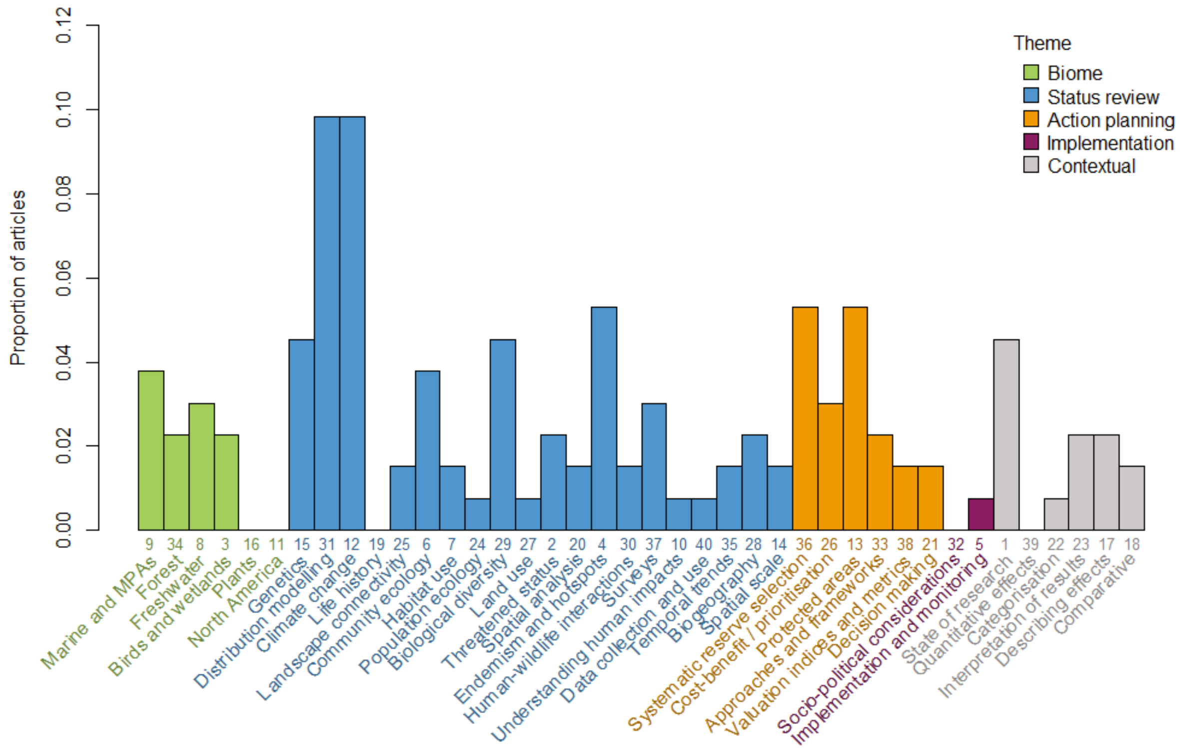


Figure A5. The distribution of documents among topics for the five journals that contributed the largest number of documents to the corpus: (a) Biological Conservation (b) Conservation Biology (c) Plos One (d) Biodiversity and Conservation, and (e) Diversity and Distributions. The proportion of documents assigned to each topic is presented in order to allow comparison among journals that have contributed different total numbers of documents to the corpus (see Fig A4).